*EVOLUTION OF IMAGE-BASED STEREO FOR SEA SURFACE NAVIGATION

Dr. Curtis Padgett

Group Supervisor Maritime and Aerial Perception Systems

Jet Propulsion Laboratory, California Institute of Technology

Mail Stop 198-235, Pasadena CA 91109

**This research was sponsored by Office of Naval Research, Code 33. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Office of Naval Research.*

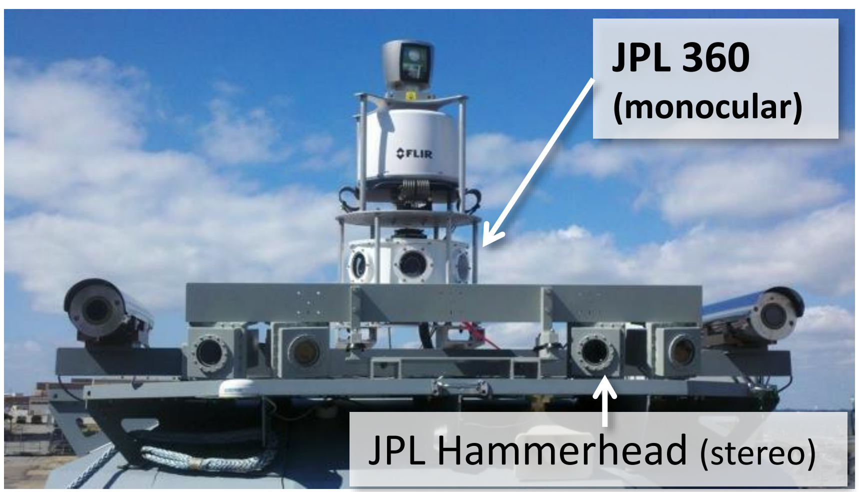
Perception Systems

JPL has extensive experience in autonomous perception, especially computer vision algorithms and software for EO/IR systems.

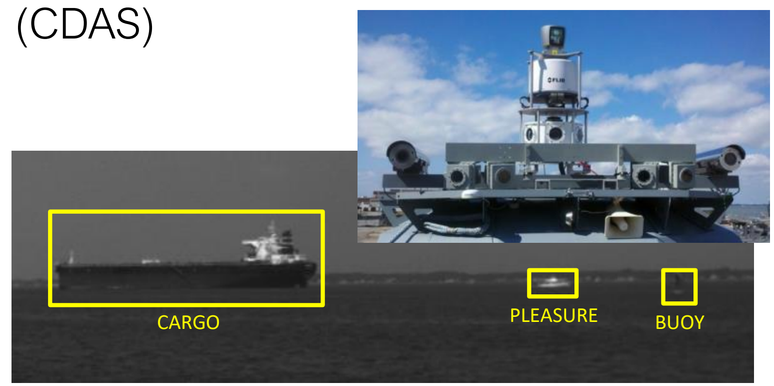
Our USV experiments often rely on solely on EO/IR perception for sitational awarness.

For USV’s we employ

- A forward-looking stereo camera system (the “Hammerhead”)
- A 360-degree monocular camera system
- A gimbaled high resolution zoom camera (EO/IR)



Contact Detection and Analysis System (CDAS)



Contact Detection and Analysis System

Jet Propulsion Laboratory
California Institute of Technology

"CDAS" is JPL's software suite for processing maritime images for USV situational awareness (monocular images, and fusion with stereo).

CDAS provides automated methods for

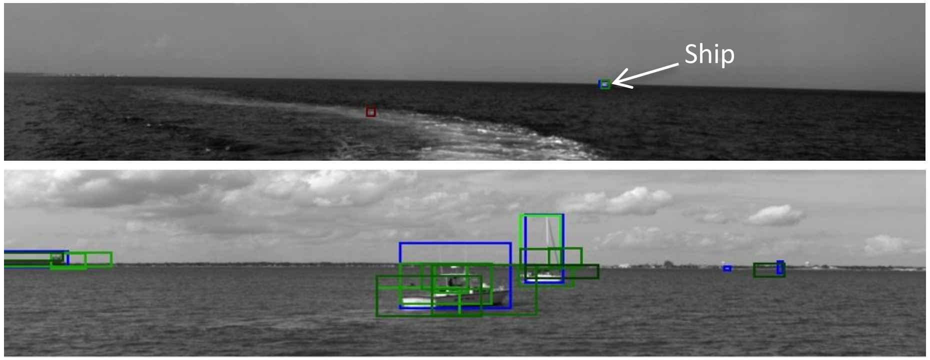
- **detection**, **tracking**, and **classification** of surface contacts

to support

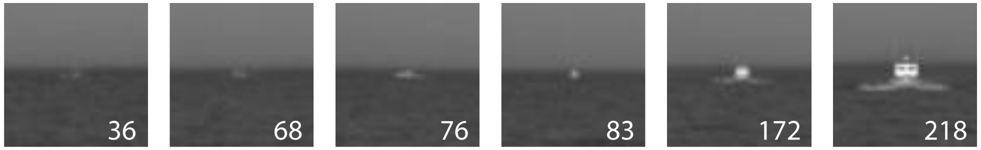
- **360-degree awareness** of other maritime traffic as specified under the International Regulations for Preventing Collisions at Sea (**COLREGS**)
- automated target recognition (**ATR**) capabilities for intelligence, surveillance, and reconnaissance (**ISR**) and other mission scenarios

Advanced detection and tracking

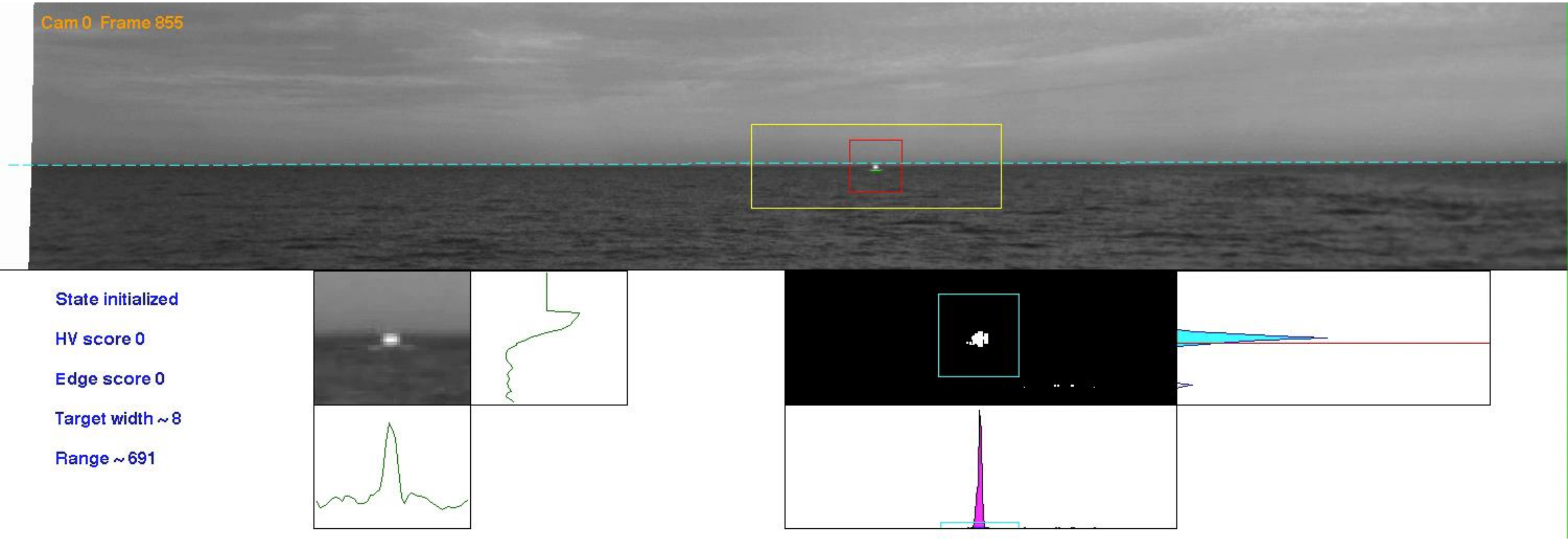
Jet Propulsion Laboratory
California Institute of Technology



Single-frame contact detection in difficult conditions, fusing up to eight custom computer vision algorithms



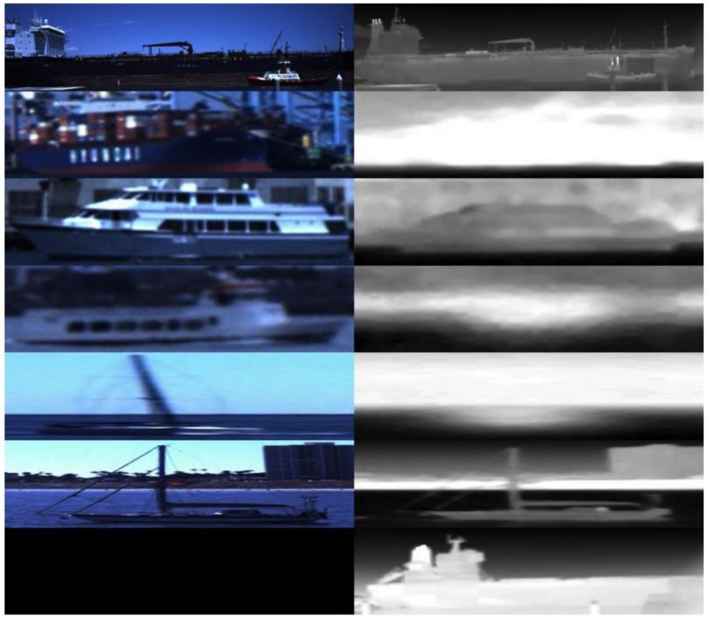
Vessel tracked over different perspectives, lighting, and scales



Vessels Are Challenging to Classify

Jet Propulsion Laboratory
California Institute of Technology

- Large changes in scale and rotation.
- Large amount of intra-class variation
  - Variations in decoration and design
  - Backgrounds
- Large amount of inter-class similarity
  - tug vs. fishing, merchant vs. military

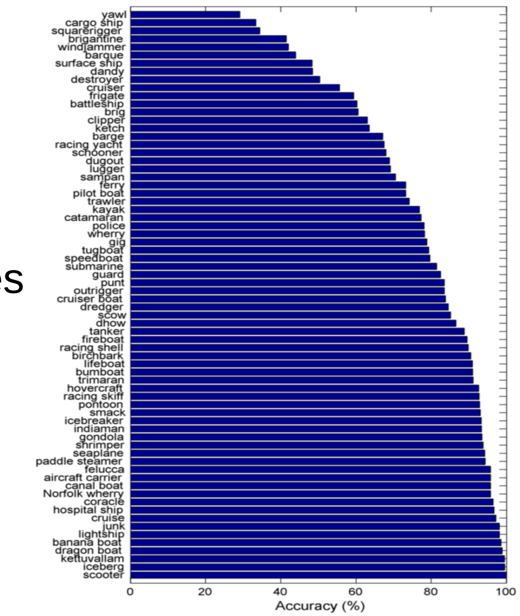


Fine-Grained Vessel Classification

Jet Propulsion Laboratory
California Institute of Technology

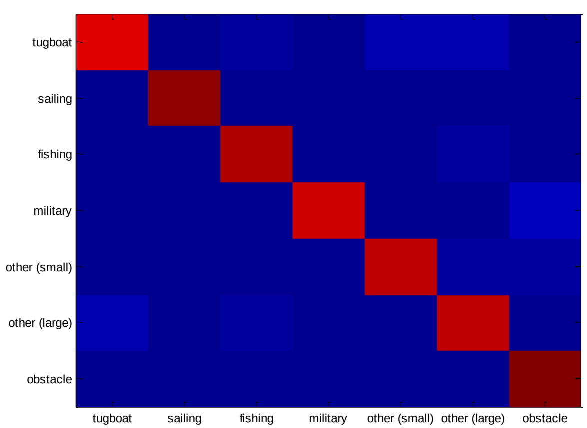
CDAS classifies vessels by type. This is needed even if for radar-centric detection, for appropriate hazard response and COLREGS compliance.

Data Overview: 70 categories, average 288 images each, of disparate examples and perspectives
Source: ImageNet



Fine-Grained classification of all 70 categories

- **Mean Per-Class Accuracy: 78.8%**
- Chance is 1.4% (1/ 70)



Salient-level classification of 7 groupings

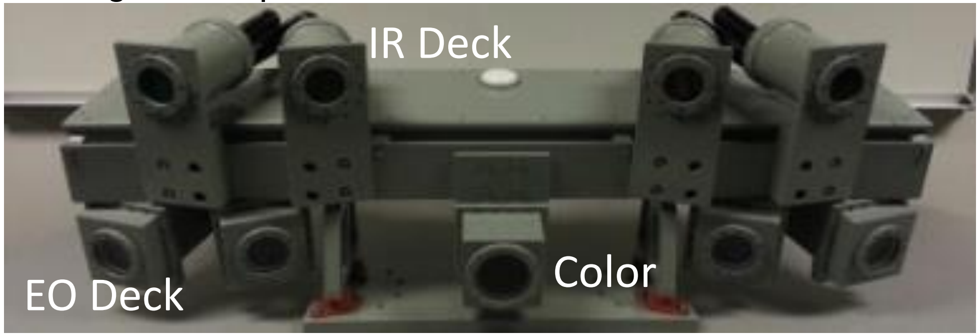
- **93.9% mean accuracy**
- These results are from fusing fine-grained results to create salient predictions.

JPL Stereo Systems

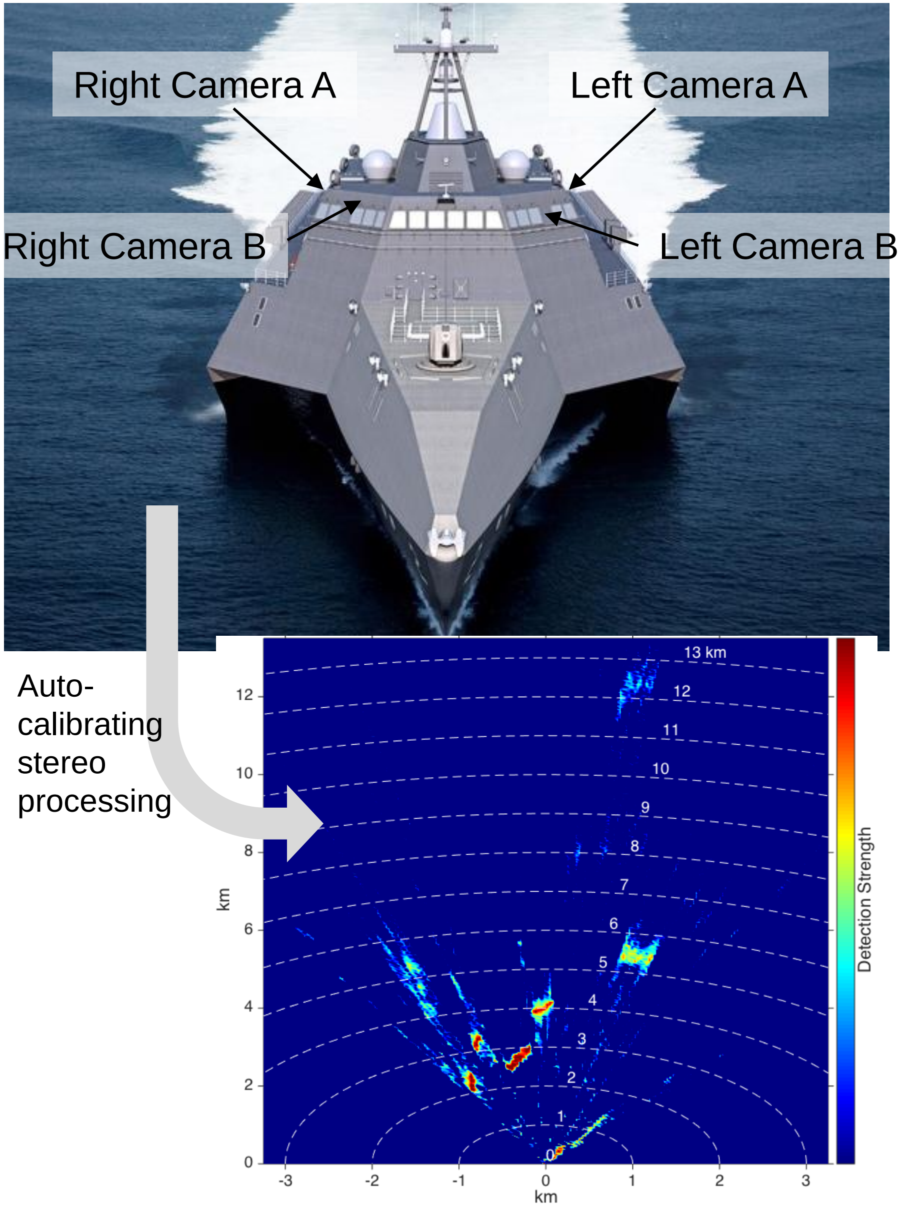
Generation 1: The Original “Hammerhead”

Generation 2: High Resolution “TTA Hammerhead”
Technology Transition Agreement (Level C), PMS 406

Generation 3: “Swampworks Hammerhead” with IR for nighttime operations



Generation 4:
Mechanically Uncoupled Stereo EO/IR (MUSE)



<u>Solution</u>:

Long-range stereo camera system for radar-free contact detection

- Detect contacts at long ranges by using the vessel as the stereo baseline
- Game-changing capability for missions requiring *passive-only sensing*
- Ship fit—just place individual cameras
- Can also map shoreline for GPS-free nav.

<u>Challenges</u>:

- No rigid stereo bar, structure will flex—needs camera calibration on every frame
- Calibration is typically a laborious manual process

<u>JPL Innovations</u>:

- Online high-fidelity auto-calibration using only operational images
- Fast stereo processing for high-resolution wide-baseline stereo systems
- Novel algorithms for learning 3D structure from multiple camera views

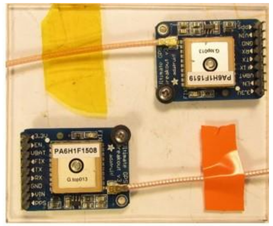
Generation 5: Cross-USV Stereo (CUS)

GPS Synchronization Boards

Two GPS modules that independently generate a pulse at the same time so that two cameras can take images at the same time.

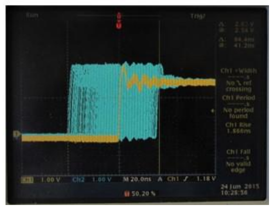
Use GPS clock for synchronization and camera pair baseline separation.

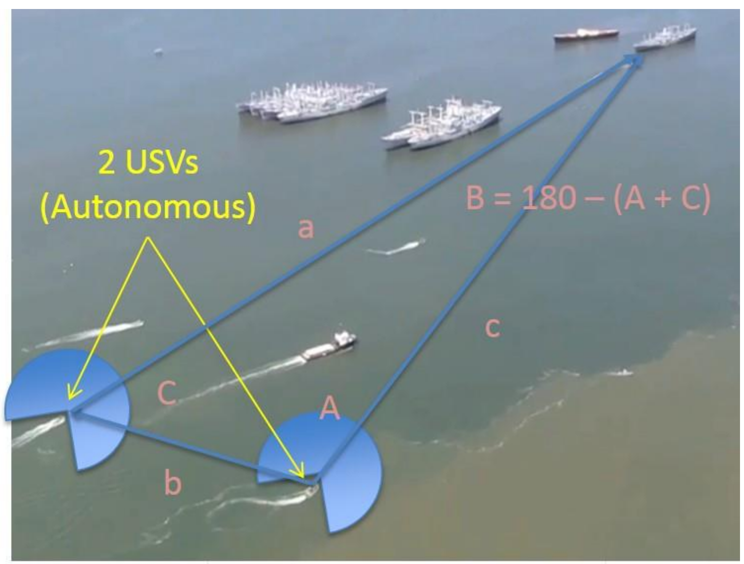
Can generate synchronized pulse if subsequently moved to a GPS denied space



Synchronization Accuracy

The above scope picture shows the spread PPS of the second GPS (cyan) relative to the first GPS (yellow). The data was collected over about 40min; the two PPS are within +/- 50ns and most of time within +/-20ns





Solution:

Long-range multi-USV triangulation

- Detect contacts at long ranges by using the separation between two vessels as the baseline
- Game-changing capability for missions requiring *passive-only sensing*
- Simultaneous capture of contacts from two vessels
- Solve triangle using angles to contact and known range between vessels

Challenges:

- Need synchronization of images separated by 100s of meters
- Find same contact at different viewpoints
- Establish precise pointing angles
- USVs work in tandem to acquire contact location

JPL Innovations:

- Use contact detection (monocular CDAS) to locate contacts at distance
- Simultaneous image capture
- Precise angle determination, multi platform

GENERATION 6
VERTICAL INTEGRATED PASSIVE EO RANGER (VIPER)

Problem:

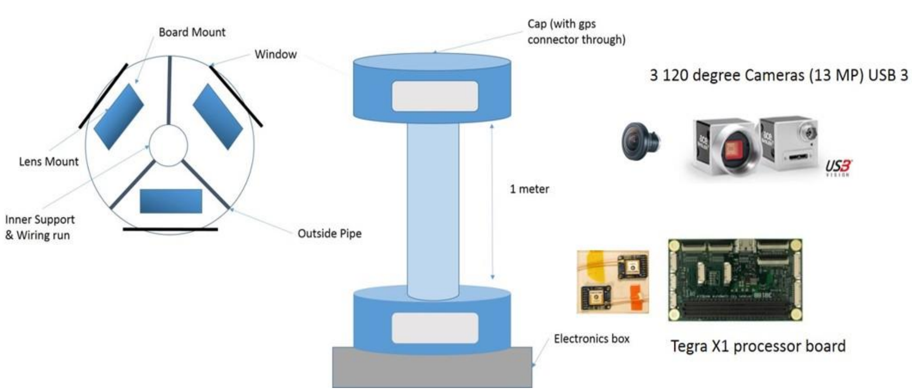
Current passive sensing solutions for SWARM or small USVs are

- heavy (> 100 lbs)
- require significant power (> 500 watts)
- relatively expensive (nearly $70K)

Solution:

A vertical stereo approach with an integrated, low power processor that uses passive, low cost EO cameras for range estimation—VIPER

VIPER Concept



Challenges:

- Cheapest cameras (< $500) are progressive scan—stereo can be degraded on a moving platform if the scan time is long
- Small (credit card size), cheap (< $1000), low power (< 5 watt) processors don’t use standard processing architecture—require software accommodation

JPL Innovations:

- Demonstrated stereo (both thermal and visible) on moving platforms using progressive scan cameras
- Demonstrated stereo code (operating 1 Hz) on Tegra X1 processor card (expect MUSE optimizations to improve performance)

VIPER Benefits:

- Power: 35 Watts continuous, < 50 Watts Peak
- Weight: < 10 lbs
- Performance: equal to or better than HammerHead (update rate ~2 Hz)
- Cost: < $10K per unit

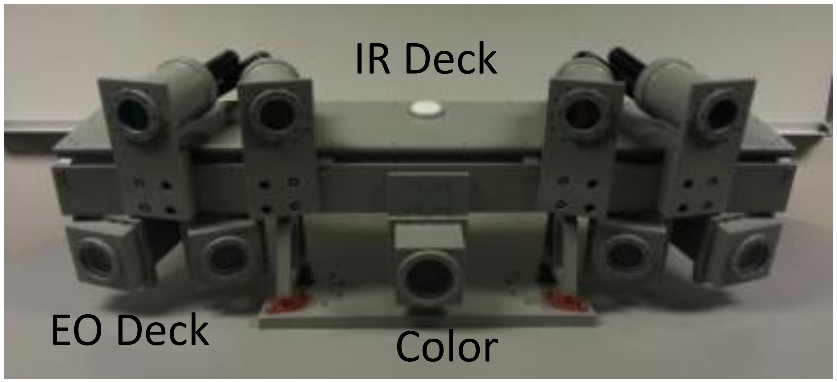
Evolution of on Water Passive Ranging

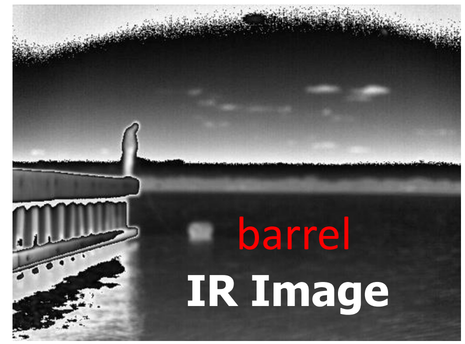
System	Max Range	EO/IR ?	Field of Regard	Platform
Hammerhead	Mid range	EO	Narrow	Powerboat
TTA Hammerhead	Long range	EO	Narrow	Powerboat
SwampHammer	Long range	EO/IR	Narrow	Powerboat
MUSE	Very long range	EO	Wide	SeaHunter
CUS	Very long range	EO	Full	USV Swarm
VIPER	Long range	EO	Full	Powerboat; Amphibious

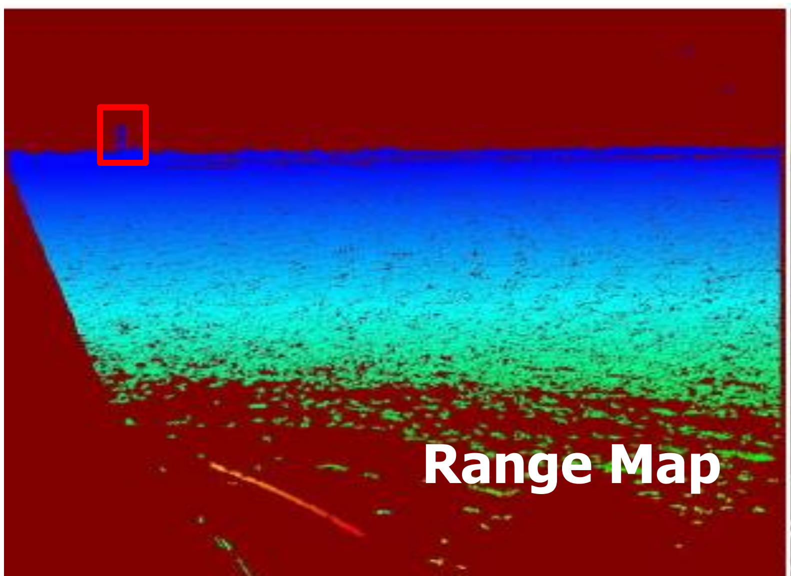
Hammerhead Stereo System

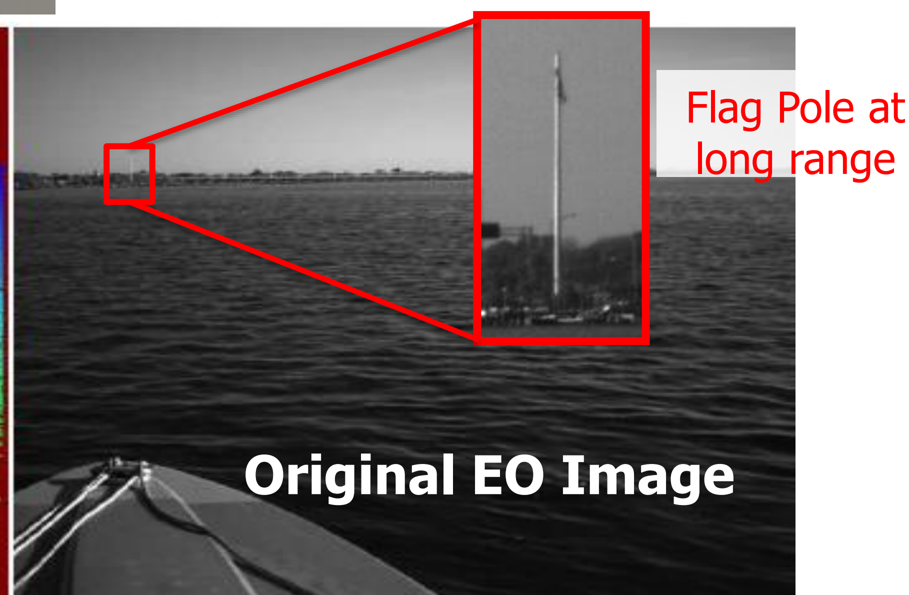
Jet Propulsion Laboratory
California Institute of Technology

- Used for forward-looking hazard detection and tracking.
- Stereo provides vision-based passive ranging, can find hazards unseed by radar.
- Hammerhead is, to our knowledge, the most capable stereo system in operation.

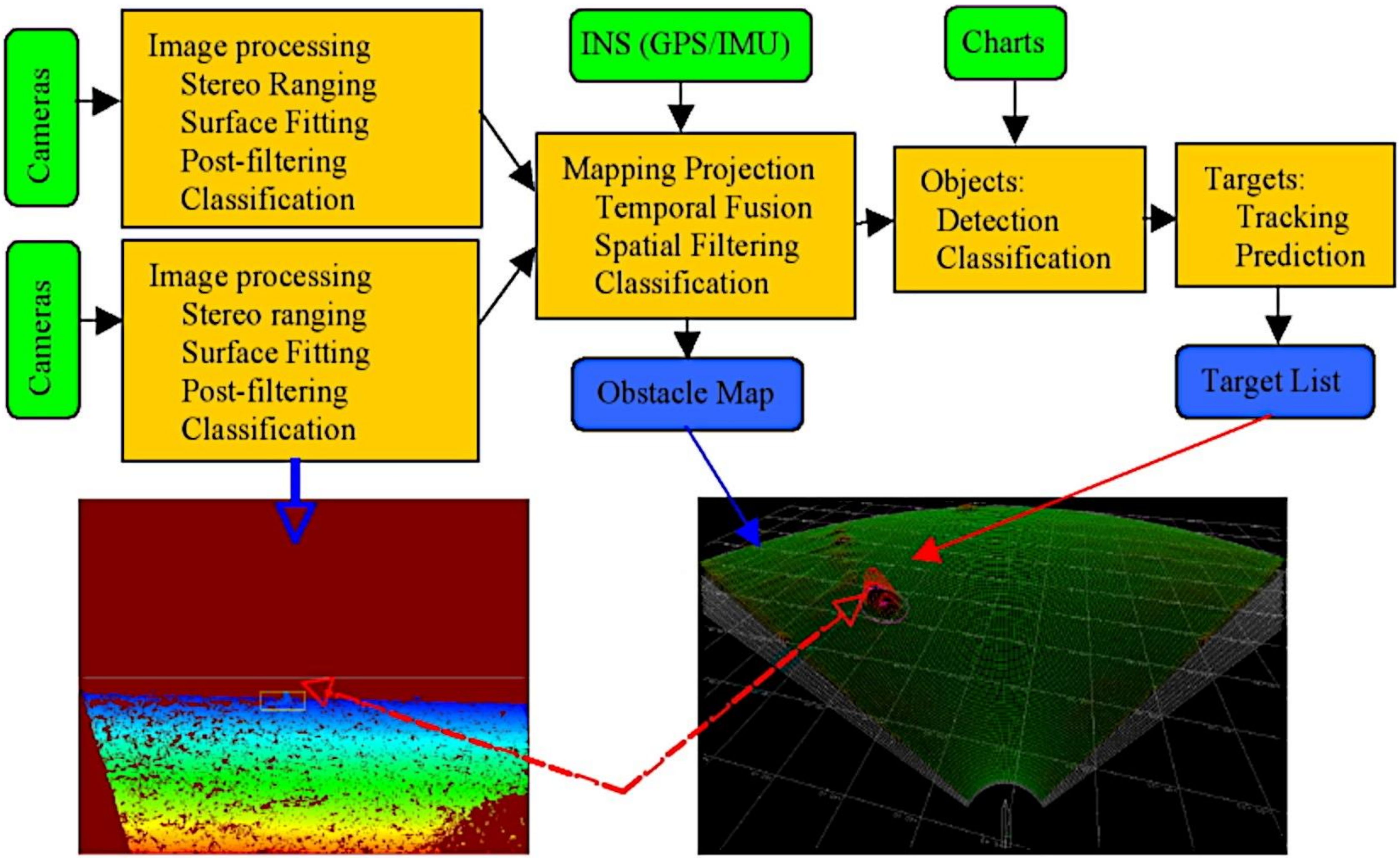






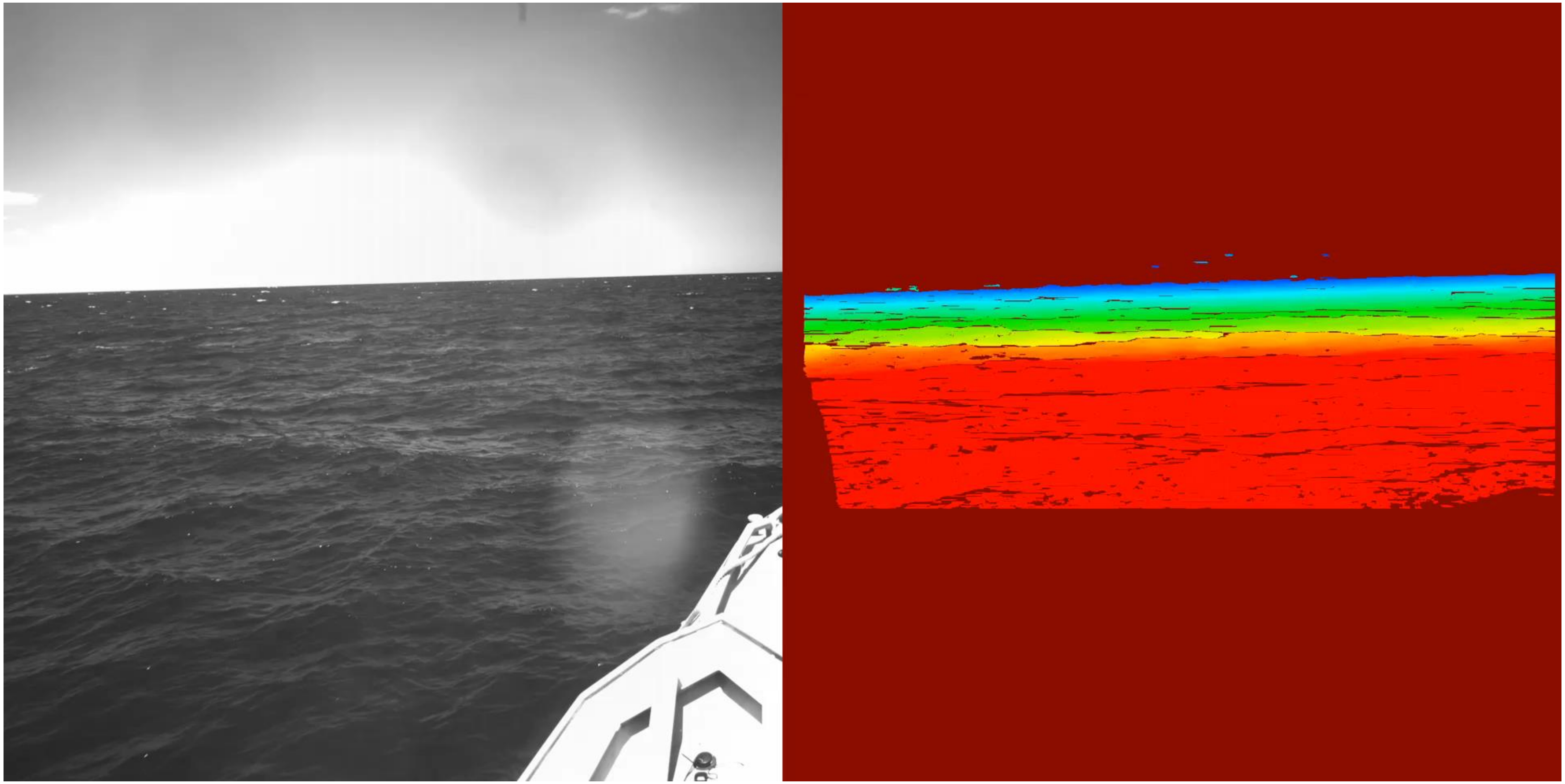


Stereo Processing Pipeline

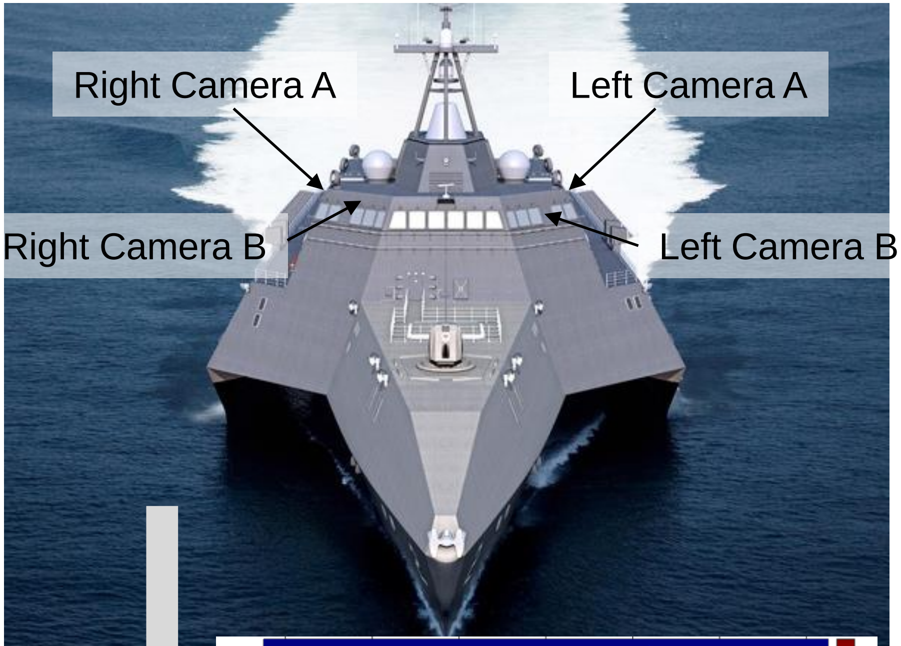


Provides Visual Detection *and Ranging*

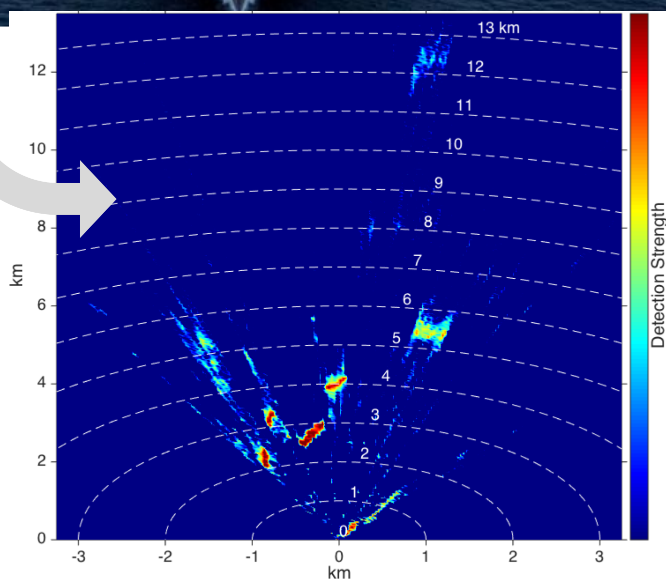
Stereo Pipeline Movie



Next Generation: Mechanically Uncoupled Stereo EO/IR (MUSE)



Auto-calibrating stereo processing



<u>Solution</u>:

Long-range stereo camera system for radar-free contact detection

- Detect contacts at long ranges by using the vessel as the stereo baseline
- Game-changing capability for missions requiring *passive-only sensing*
- Ship fit—just place individual cameras
- Can also map shoreline for GPS-free nav.

<u>Challenges</u>:

- No rigid stereo bar, structure will flex—needs camera calibration on every frame
- Calibration is typically a laborious manual process

<u>JPL Innovations</u>:

- Online high-fidelity auto-calibration using only operational images
- Fast stereo processing for high-resolution wide-baseline stereo systems
- Novel algorithms for learning 3D structure from multiple camera views

MUSE System for Long Range Situational Awareness

Jet Propulsion Laboratory
California Institute of Technology

- Used for forward-looking hazard/contact detection and tracking.
- Stereo provides image-based passive ranging.
- To be deployed on Sea Hunter for testing in May.

MUSE Challenges

Jet Propulsion Laboratory
California Institute of Technology

- **Large Baseline**
  - Totally rigid mounting bar impractical at long distance (>> 1 meter)
  - Weight issues (very heavy bar to get needed rigidity)
  - Accommodation (large, heavy things with large dimension and field of view requirements make placement very difficult)
  - Perspective distortion—large displacement between imagers means that the near field doesn't look the same between left and right pair
- **High Resolution Imagers**
  - Need very large imagers to "see" contacts
  - Need small angular resolution to get meaningful disparity (range estimates)
  - Requires high precision camera models for fast stereo implementations
  - Means large processing bandwidth or very slow update rates (not practical on high speed platform)
- **Wide coverage (field of view)**
  - Need to see all contacts in front of vehicle to provide reasonable situational awareness
  - COLREGs requires that we know about crossers and respond appropriately on open water (large stereo typical moves at long distances)

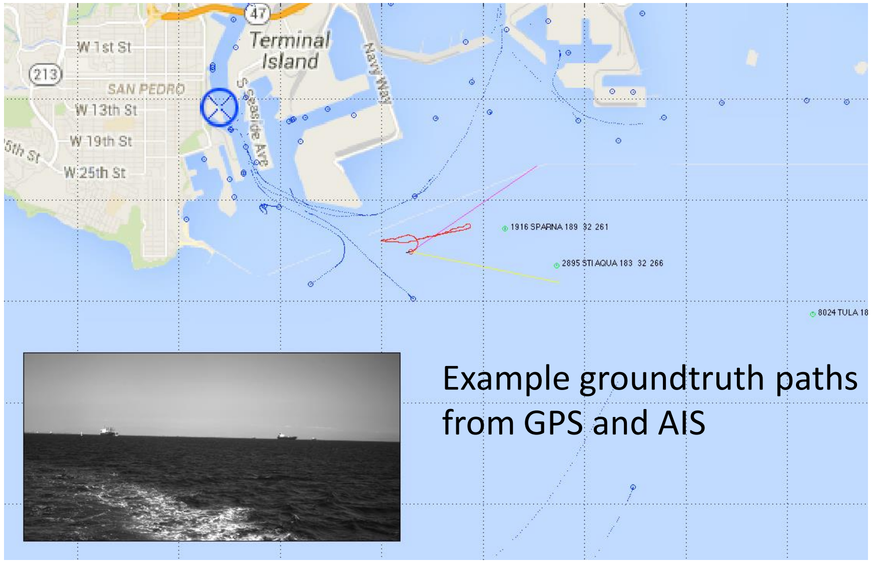
MUSE Technical Approach

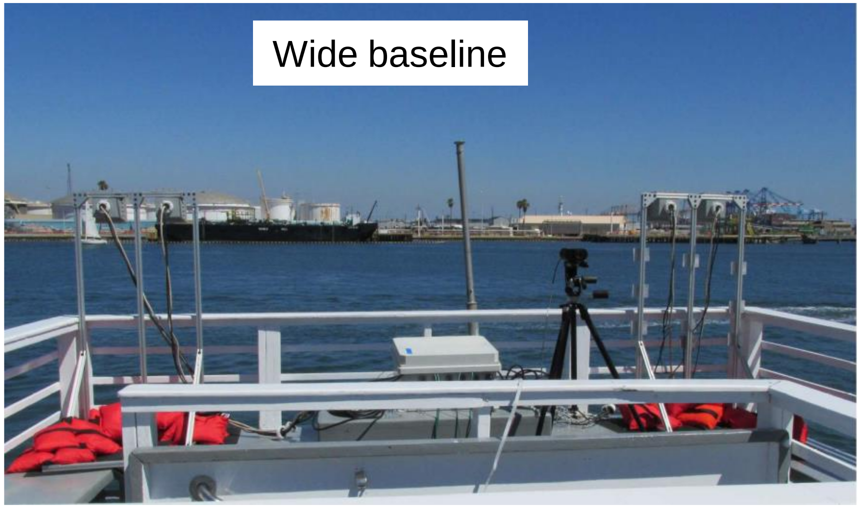
- **Mount cameras independently**
  - Constrained by field of view considerations—do not require rigid, coupled mounting
- **Camera relative motion**
  - "Solve" extrinsic camera parameters for each stereo frame to take out any relative motion
  - Use knowledge of sea surface to remove perspective changes between images
  - Search for common features—use to precisely determine relative pose and sea surface
- **Improve stereo processing**
  - Parallelize everything to improve throughput

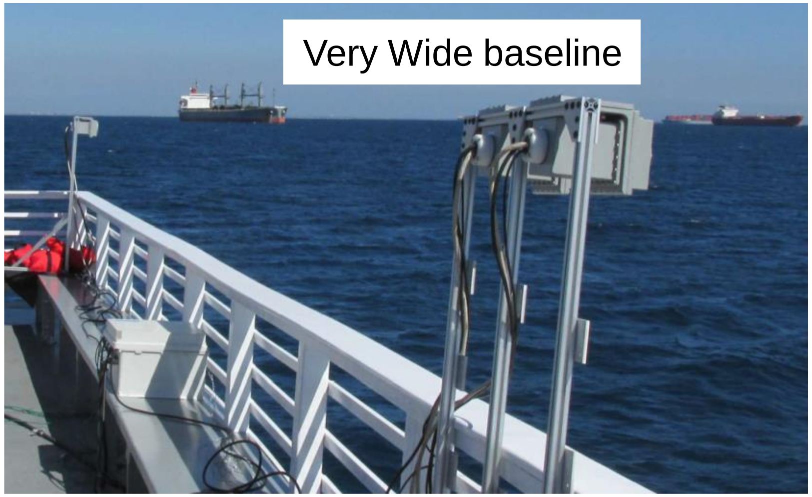
On-Water Data Collection and Testing

Jet Propulsion Laboratory
California Institute of Technology

- Chartered boat in LA Harbor, 9/17/15
- Mounted 2 stereo camera pairs
  - “Wide” baseline
  - “Very Wide” baseline
- Ground Truth: AIS, radar, range gun

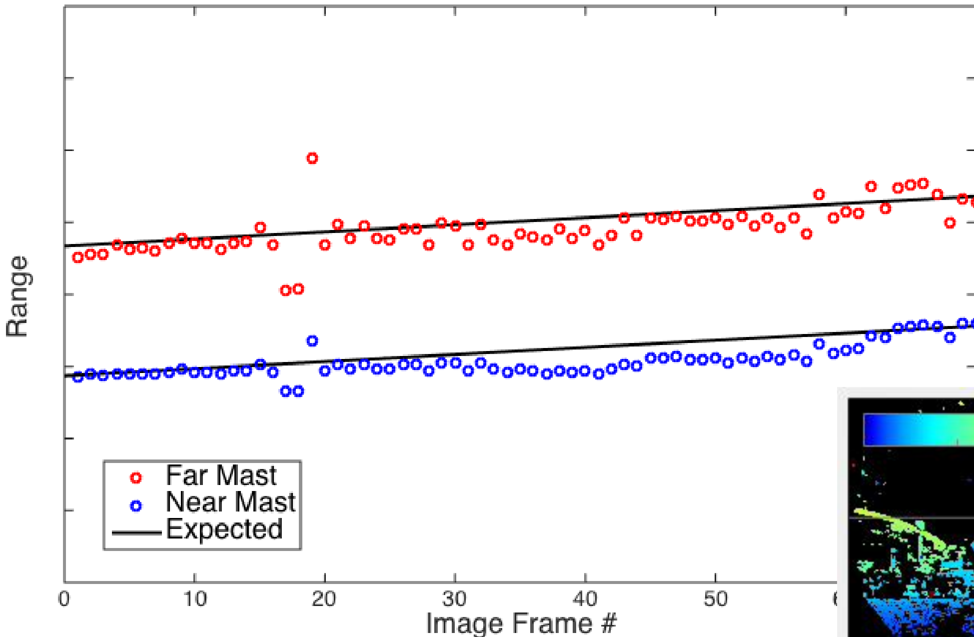






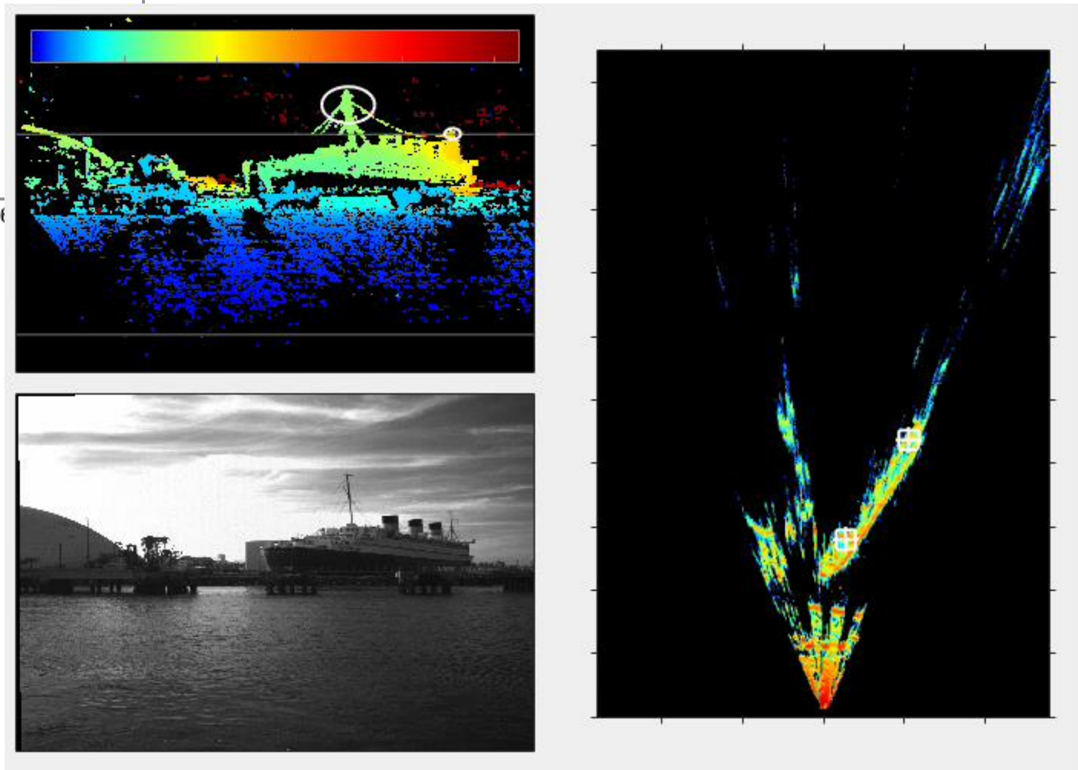
Calibration Test: Queen Mary

Jet Propulsion Laboratory
California Institute of Technology



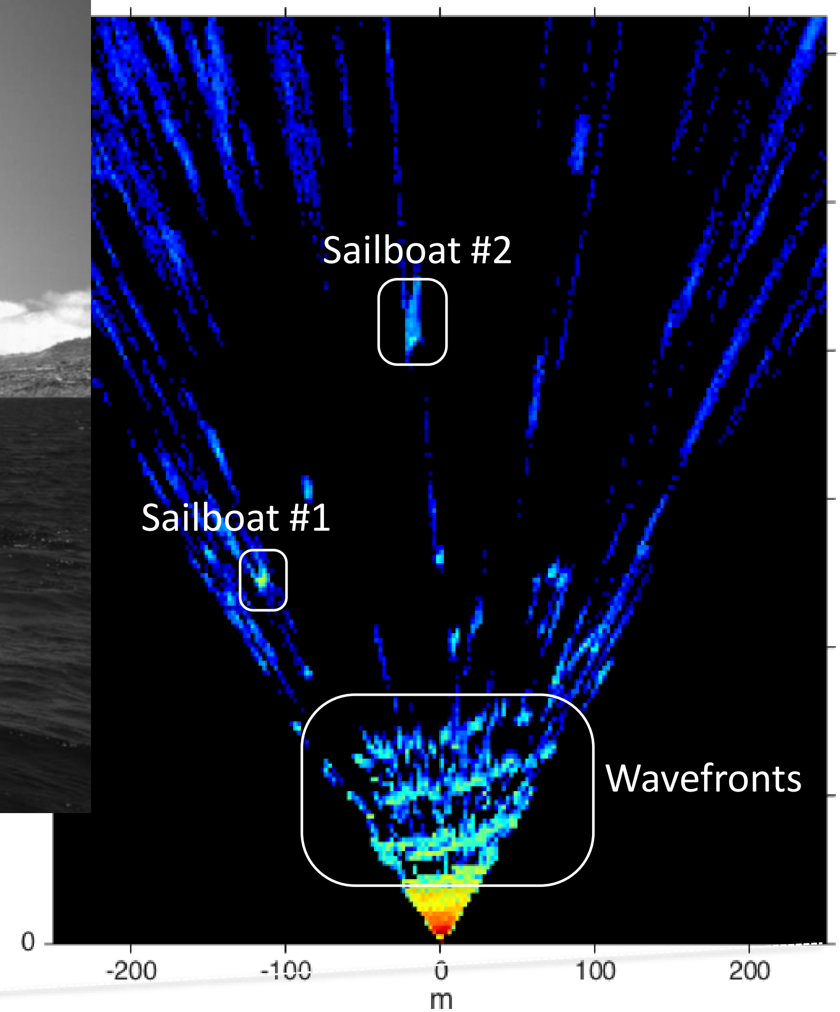
Stereo range to both masts over a continuous period with frame-wise auto-calibration

Expected ranges based on current GPS position



MUSE: Example Imagery/Ranging

Jet Propulsion Laboratory
California Institute of Technology

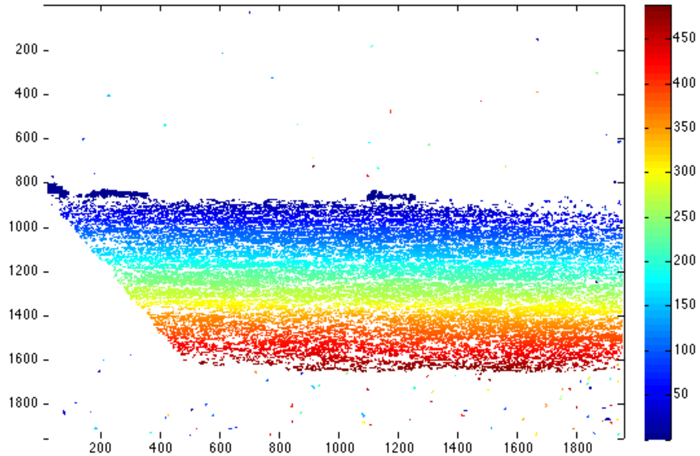


Stereo Processing

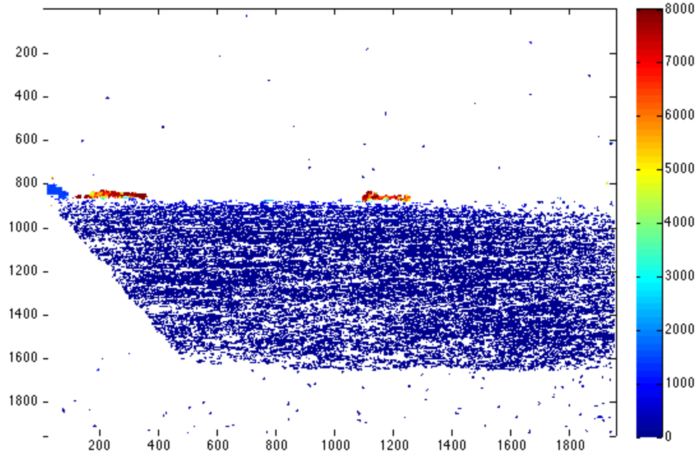
Jet Propulsion Laboratory
California Institute of Technology

Fiesta data 3 m baseline – Sample offline stereo result using Deltapose to adjust right camera model

Disparity map (pixels)



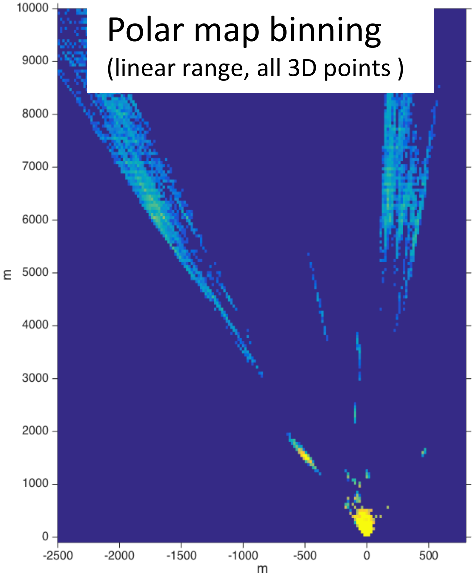
Range map (m)

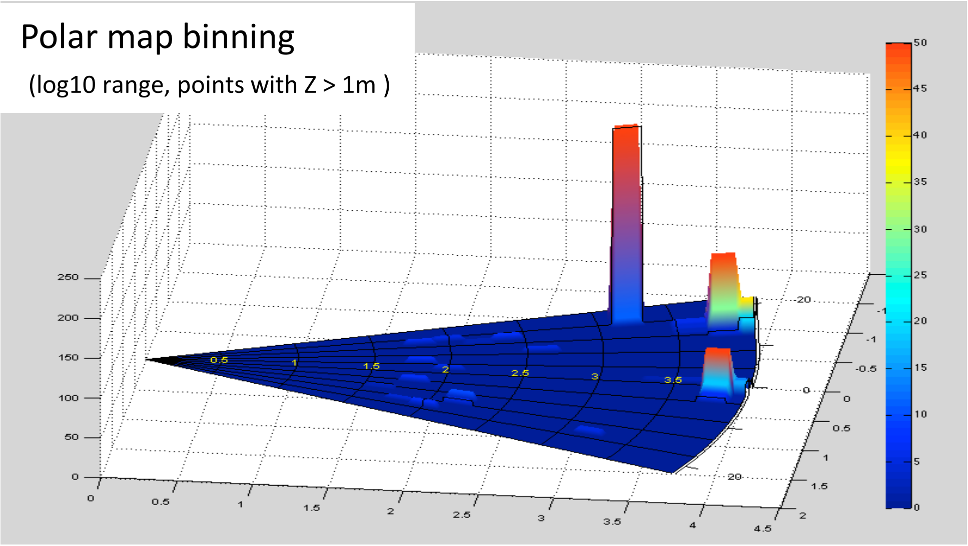


Detection Processing

Jet Propulsion Laboratory
California Institute of Technology

Estimate water plane from 3D point cloud, rotate into XY, and bin





Cross-USV Stereo (CUS)

GPS Synchronization Boards

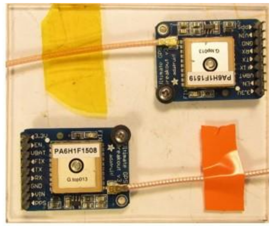
Two GPS modules that independently generate a pulse at the same time so that two cameras can take images at the same time.

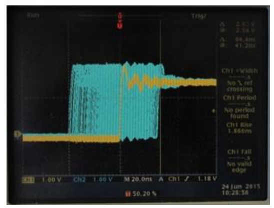
Use GPS clock for synchronization and camera pair baseline separation.

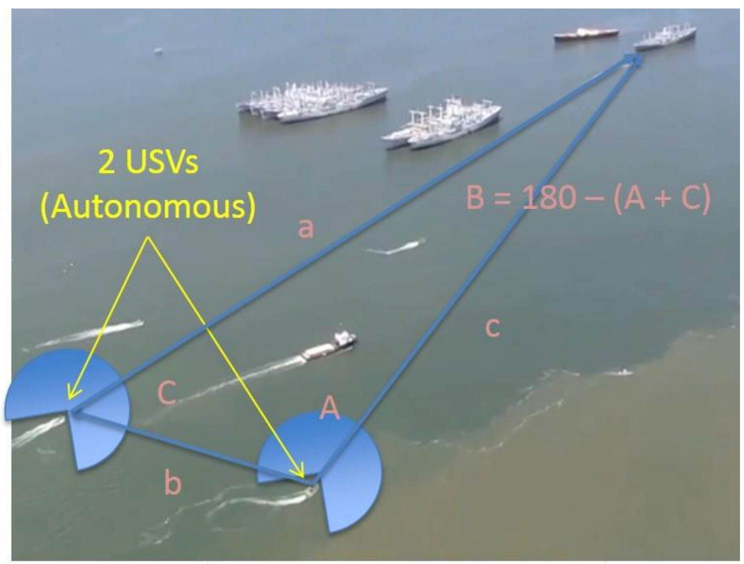
Can generate synchronized pulse if subsequently moved to a GPS denied space

Synchronization Accuracy

The above scope picture shows the spread PPS of the second GPS (cyan) relative to the first GPS (yellow). The data was collected over about 40min; the two PPS are within +/- 50ns and most of time within +/-20ns







<u>Solution</u>:

Long-range multi-USV triangulation

- Detect contacts at long ranges by using the separation between two vessels as the baseline
- Game-changing capability for missions requiring *passive-only sensing*
- Simultaneous capture of contacts from two vessels
- Solve triangle using angles to contact and known range between vessels

<u>Challenges</u>:

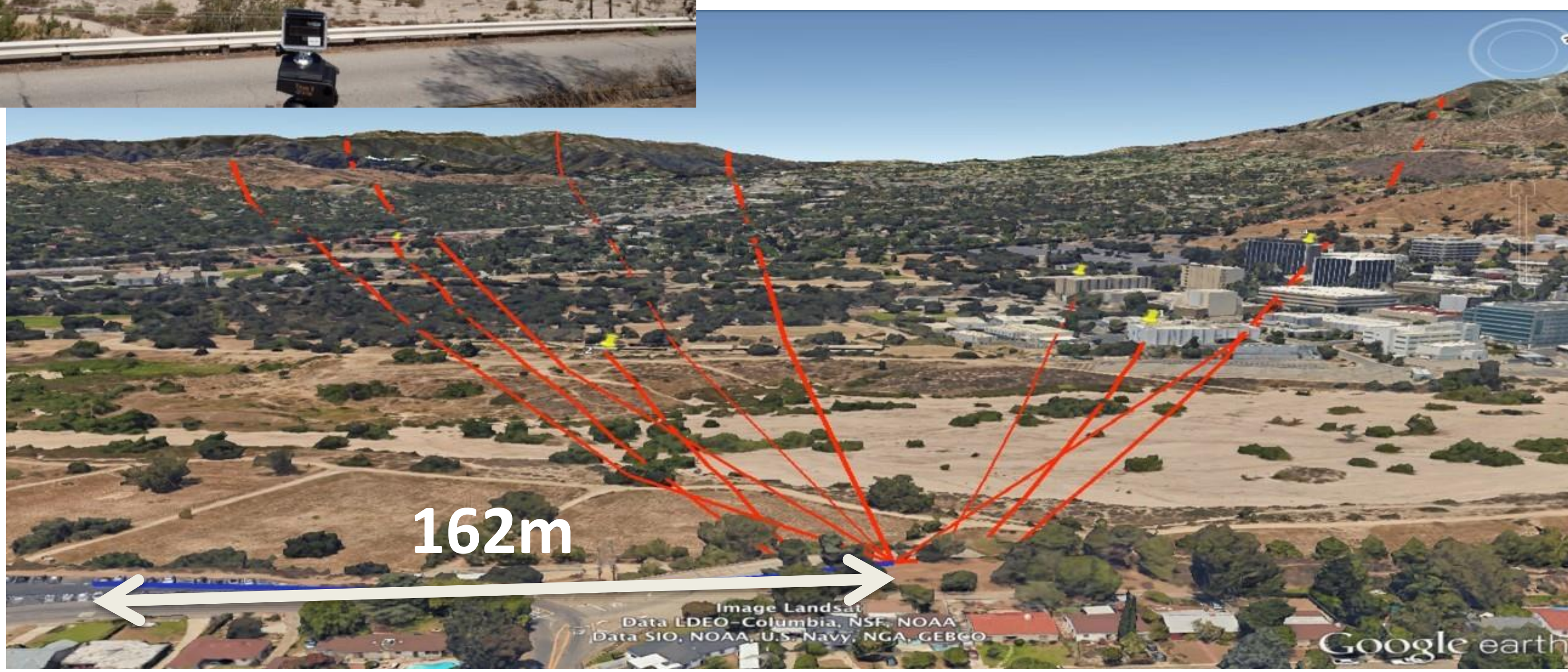
- Need synchronization of images separated by 100s of meters
- Find same contact at different viewpoints
- Establish precise pointing angles
- USVs work in tandem to acquire contact location

<u>JPL Innovations</u>:

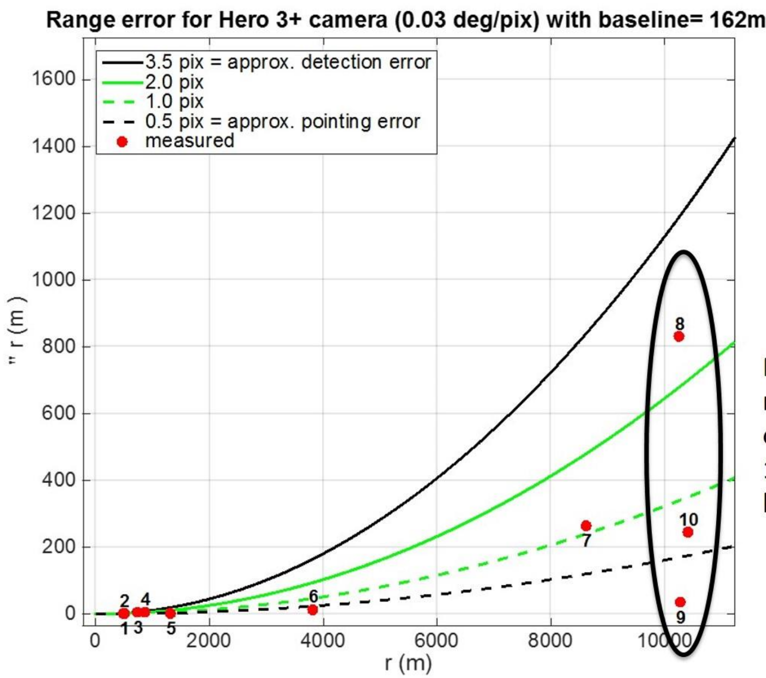
- Use contact detection (monocular CDAS) to locate contacts at distance
- Simultaneous image capture
- Precise angle determination, multi platform

CUS Test Set Up with 162m baseline (1 of 2)

Jet Propulsion Laboratory
California Institute of Technology

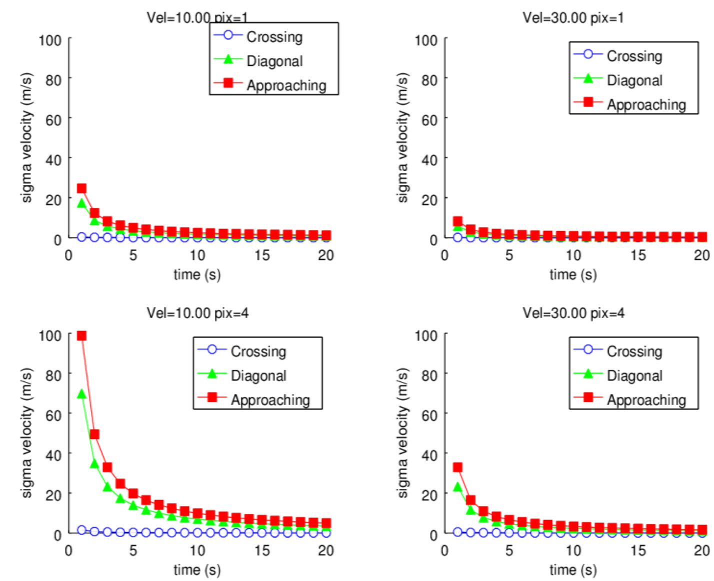


CUS Results



Mean range error of targets at 10.1-10.4 km:

3.5 %



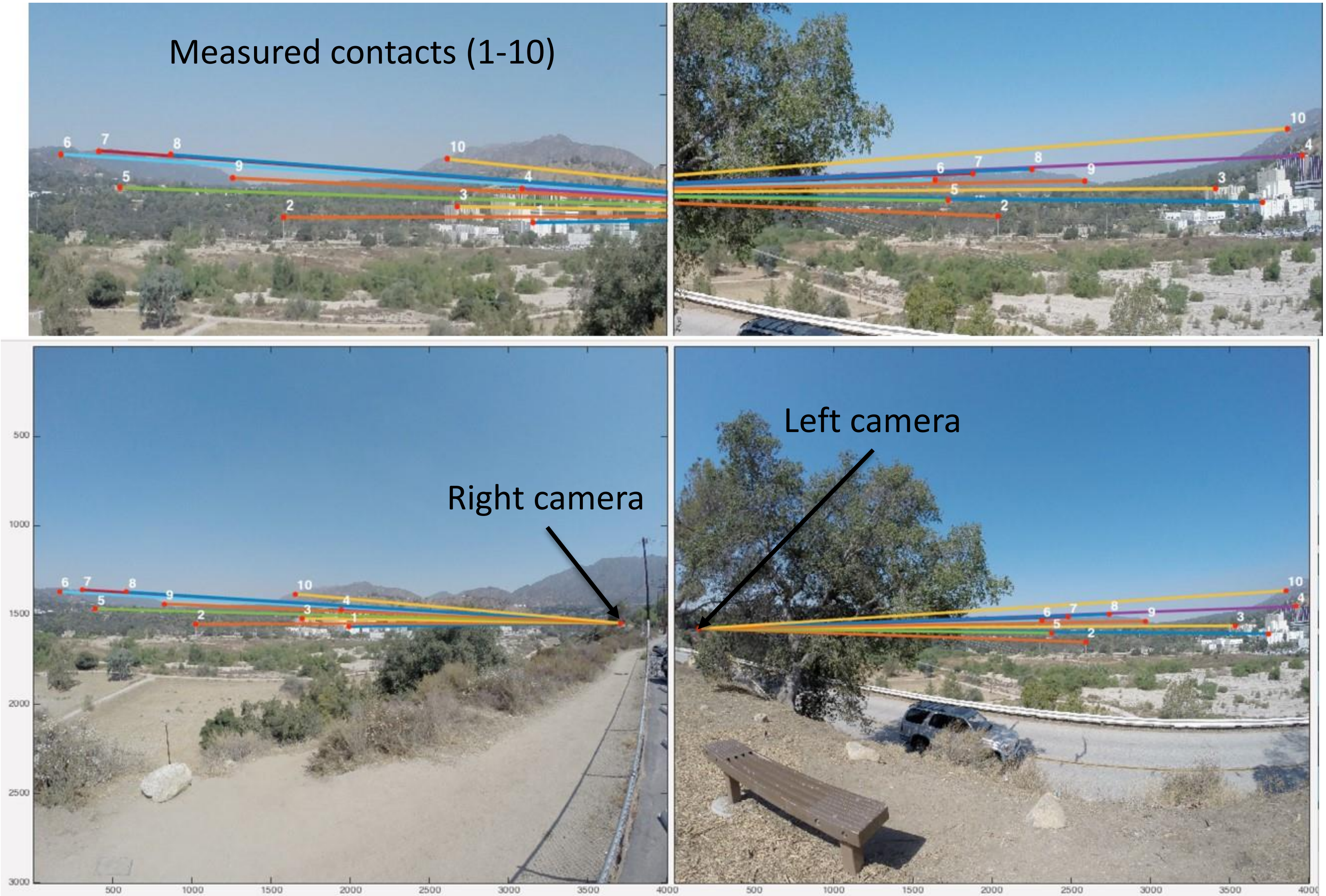
Error as a function of range for one observation with 162 meter separation between cameras.

Observed error approximately 1 pixel in magnitude (3.5% with this baseline @ 10 km, 350 meters)

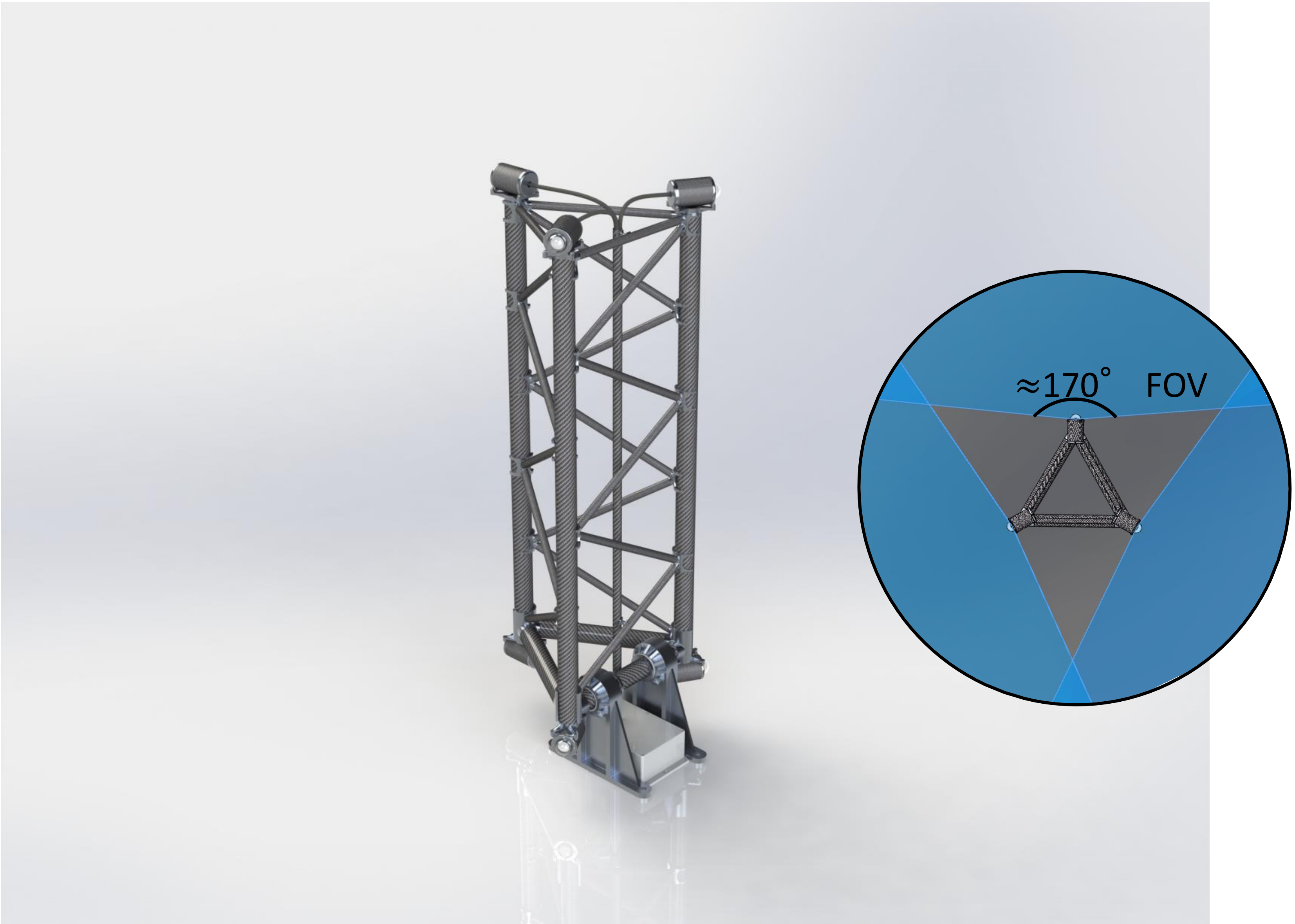
Velocity error over number of observations for given directions and pixel error @ 10 km.

Take home: Velocity error under 50% after 5 observations with (expected) 1 pixel error for worst case direction (coming right at us). In a SWARM, this could be accomplished with all USVs providing bearing to contact (first observation).

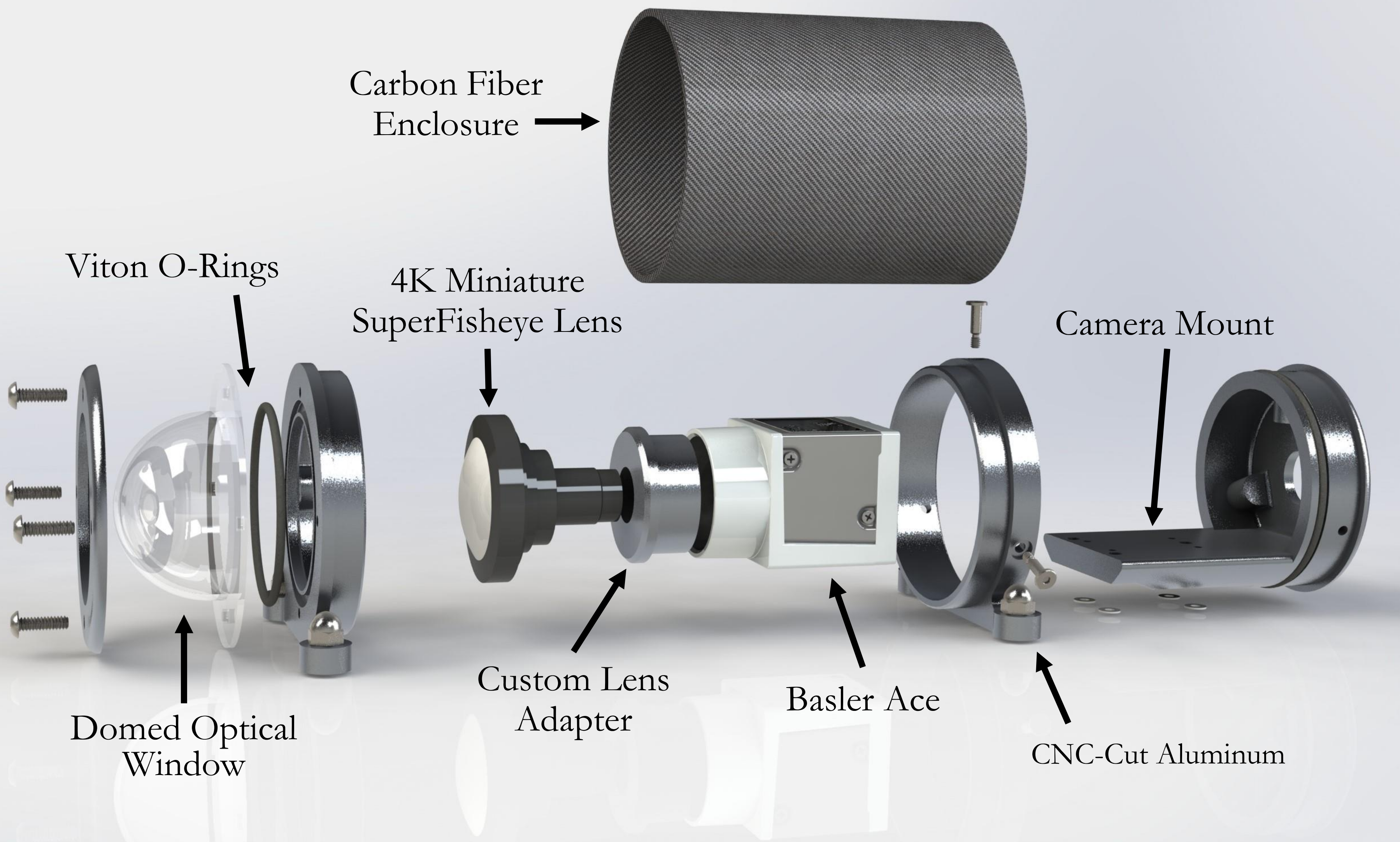
Test Setup 162 m baseline (2 of 2)



Mechanical Overview
CUS/VIPER



Camera Module



Summary

JPL has developed a number of sensor technologies to enable autonomous surface craft navigation and situational awareness.

- Expanded range and accuracy
- Solved difficult calibration issues
- Provided hazard detection and tracking in multiple modalities
- Visual classification of vessels
- Extending stereo for long range applications